

GenAD: Generative End-to-End Autonomous Driving

Year: 2024 | Citations: 129 | Authors: Wenzhao Zheng, Ruiqi Song, Xianda Guo, Long Chen

Abstract

Directly producing planning results from raw sensors has been a long-desired solution for autonomous driving and has attracted increasing attention recently. Most existing end-to-end autonomous driving methods factorize this problem into perception, motion prediction, and planning. However, we argue that the conventional progressive pipeline still cannot comprehensively model the entire traffic evolution process, e.g., the future interaction between the ego car and other traffic participants and the structural trajectory prior. In this paper, we explore a new paradigm for end-to-end autonomous driving, where the key is to predict how the ego car and the surroundings evolve given past scenes. We propose GenAD, a generative framework that casts autonomous driving into a generative modeling problem. We propose an instance-centric scene tokenizer that first transforms the surrounding scenes into map-aware instance tokens. We then employ a variational autoencoder to learn the future trajectory distribution in a structural latent space for trajectory prior modeling. We further adopt a temporal model to capture the agent and ego movements in the latent space to generate more effective future trajectories. GenAD finally simultaneously performs motion prediction and planning by sampling distributions in the learned structural latent space conditioned on the instance tokens and using the learned temporal model to generate futures. Extensive experiments on the widely used nuScenes benchmark show that the proposed GenAD achieves state-of-the-art performance on vision-centric end-to-end autonomous driving with high efficiency. Code: <https://github.com/wzzheng/GenAD>.